



Linking IATE, EUROVOC and Euramis

Session 7 – The Luxembourg BabelNet Workshop [13:15-14:45, 3 March, 2016]

- We will see **different strategies** to link three different resources to BabelNet (and, as a result, to each other):
 - IATE
 - EUROVOC
 - Euramis
- More resources can be linked with similar strategies, so it is just up to you to decide what to work on!
- You could add these resources to the LLOD!



We have a dream!







- **IATE** (Inter-Active Terminology for Europe) is the EU's interinstitutional terminology database.
- Administered by the EU Translation Centre
- Legacy databases imported into IATE:
 - Eurodicautom
 - o TIS
 - Euterpe
 - Euroterms
 - CDCTERM
- Since 2002 it has been enriched by all EU translators
- A public version contains 8.5M validated terms from EU legislation in 110 domains in 24 languages





English (en)

Search Screen Help

marketization

en > Any (domain: Any domain, type of search: All)

Result 1- 1 of 1 for marketization

FINA	ANCE [COM]	<u>Full entry</u>	
EN	marketization	*** *@ <u></u>	
DA	marketisation	*** *@	
DE	Marktbildung in standardisierter Form	**** * @	
EL	εμπορευματοποίηση	*** *@	
ES	mercadización	*** *@	
FI	markkinavetoisuus	**** *@	
гі	markkinalähtöisyys	** ** * @	
FR	marchéisation	*** * @	
IT	marketization	*** *@	
NL	marktgericht zijn	*** *@	
sv	marknadsorientering	*** *@ 📝 🛄	

Search

1

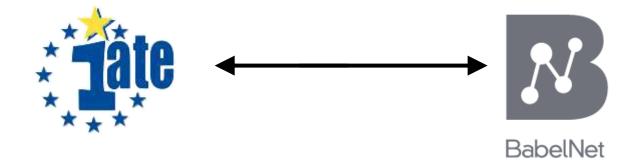
en > Any marketization

Search



Linking IATE to BabelNet

Goal: To automatically (and semantically) link IATE to BabelNet using a **language- and resource-agnostic approach**





Linking IATE to BabelNet

Consider the entries for the Italian Republic in IATE and in BabelNet





Linking IATE to BabelNet





Downloading IATE

IATE's dump is available for download at:

http://iate.europa.eu/tbxPageDownload.do

The dump contains: **lemmas**, **translations**, **domains**



Downloading IATE

The dump is missing some important information.

Here is a query on the IATE website (http://iate.europa.eu) for the English term "national competitiveness board"

not in the dump in the dump

ominio	Costruzione europea, VITA ECONOMICA, Dominio sconosciuto	
	en —	
Definizione	board in the euro area Member State set up to monitor performance and policies in the field of competitiveness, thereby contributing to foster sustained economic convergence, and to increase ownership of the necessary reforms at the national level	
Rif. della definizione	Recommendation for a Council recommendation on the establishment of National Competitiveness Boards within the Euro Area, COM/2015/0601 final, <u>52015DC0601:en:NOT</u> " > 52015DC0601/en	
	COM-Internal document: SG/2015/80295/03/00, Council Recommendation on the establishment of National Competitiveness Boards within the Euro Area [21.10.2015]	
Termine	national competitiveness board	
Affidabilità	2 (Affidabilità minima)	
Riferimento relativo ad un termine	Recommendation for a Council recommendation on the establishment of National Competitiveness Boards within the Euro Area, COM/2015/0601 final, 52015DC0601:en:NOT">52015DC0601:en:NOT > 52015DC0601/en	
	COM-Internal document: SG/2015/80295/03/00, Council Recommendation on the establishment of National Competitiveness Boards within the Euro Area [21.10.2015]	
Contesto	'The setting up of national competitiveness boards in charge of tracking performance and policies in the field of competitiveness should contribute to the enhancement of ownership of the necessary policies and reforms at national levels and to improving the knowledge basis for Union economic policy coordination in the field of competitiveness.'	
Rif. del contesto	Recommendation for a Council recommendation on the establishment of National Competitiveness Boards within the Euro Area, COM/2015/0601 final, recital 3, 52015DC0601:en:NOT" >52015DC0601/en	
Data	11/12/2015	



How to link IATE to BabelNet? (simple mapping)

By solely relying on domains, terms and translations, a first IATE-to-BabelNet mapping approach can be carried out in three steps:

- 1. Retrieve all **concepts** having the same English lemma in both resources (warning: an ambigous lemma can cover different concepts);
- 2. Compare each retrieved IATE concept against each BabelNet concept for the lemma;
- 3. Given a IATE concept, create a mapping to the BabelNet concept that **maximizes the overlap** of lemmas and translations.





How to link IATE to BabelNet? (simple mapping)



id="IATE-134173"

EN: Italy, Italian Republic

FR: l'Italie, République italienne

IT: Italia, Repubblica italiana

11.[

EN: Italy, Italian Republic, Republic of Italy ...

FR: Italie, République italienne, État italien ...

IT: Italia, Repubblica italiana, Repubblica d'Italia

id="bn:00047705n"

id="bn:00047705n"

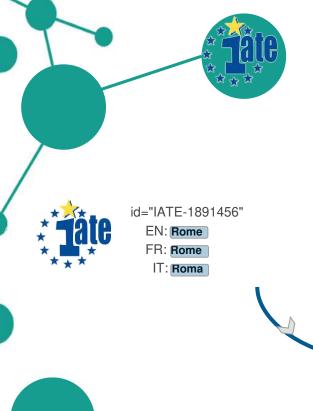
EN: Italy Texas, Italy, ...

FR: Italy



IATE-134173

bn:00047705n



How to link IATE to BabelNet? (simple mapping)

id="bn:01341486n"

EN: Rome FR: Rome

IT: Roma

id="bn:00015556n"

EN: Rome

FR: Rome

IT: Roma





We can leverage **Babelfy**, our joint graph-based approach to multilingual **Entity Linking** and **Word Sense Disambiguation**











We can leverage **Babelfy**, our joint graph-based approach to multilingual **Entity Linking** and **Word Sense Disambiguation**















Collect all the translations and synonyms and generate a multilingual input text for **Babelfy**:





Европейски парламент, Evropský parlament, Europa-Parlamentet, Europäisches Parlament, Eupωπαϊκό Κοινοβούλιο, EK, EP, European Parlament, PE, Parlamento Europeo, Europa Parlament, EP, Europan parlamentti, EP, PE, Parlament européen, PE, Parlamint na hEorpa, EP, Europski parlament, EP, Europai Parlament, Parlamentum Europæ, Parlamentum Europæum, Europos Parlamentas, EP, Eiropas Parlaments, EP, PE, Parlament Europew, EP, Europeas Parlament, PE, Parlament Europeiski, PE, Parlamento Europeu, PE, Parlamentul European, Europe

Enable partial matches:

AGNOSTIC

BABELFY!

OLOG IN REGISTER

"language agnostic" setting



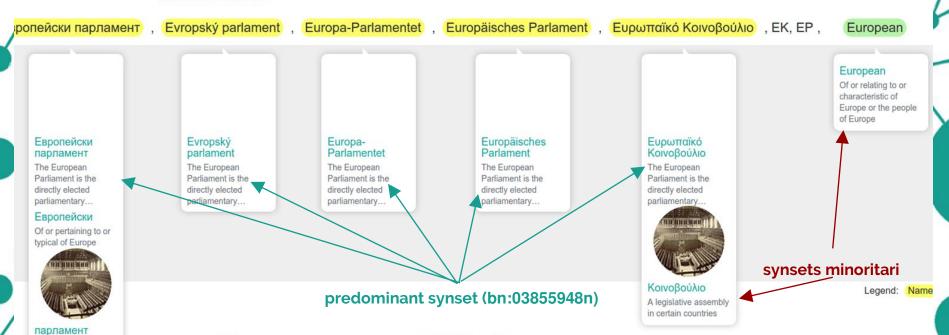
Европейски парламент, Evropský parlament. Europa-Parlamentet. Europäisches Parlament, Euρωπαϊκό Κοινοβούλιο, EK, EP, European Parliament, PE, Parlamento Europeo, Europa Parlament. EP. Europan parlamentti, EP, PE, Parlament européen, PE, Parlamint na hEorpa, EP, Europski parlament, EP, Europai Parlament, Parlamentum Europæ, Parlamentum Europæum, Europos Parlamentas, EP, Eiropas Parlaments, EP, PE, Parlament Ewropew, EP, Europees Parlament, PE, Parlament Europeiski, PE, Parlamento Europeu, PE, Parlamentul European, Europ

Enable partial matches:

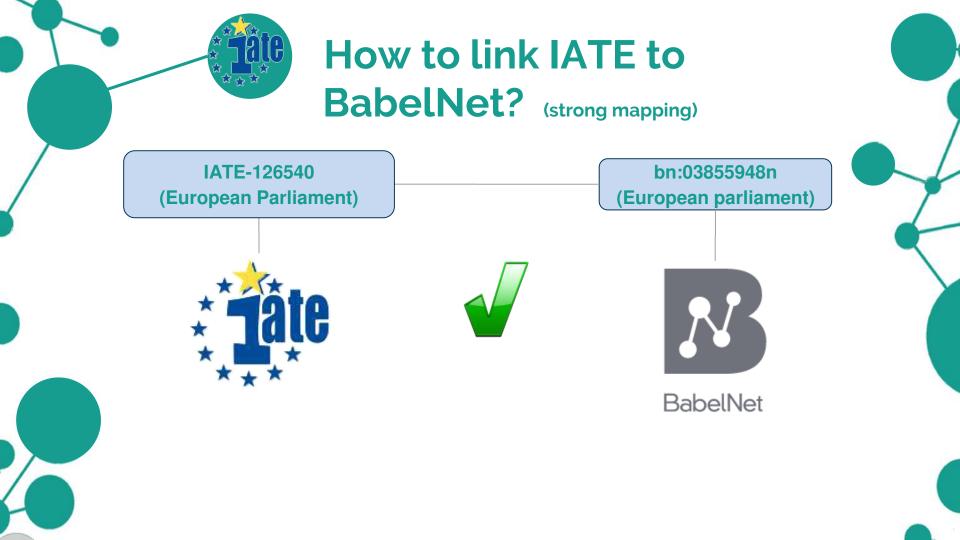
AGNOSTIC

BABELFY!

expanded view | compact view



Bahelfy is an output of the Multi IEDI ERC Starting Grant No. 259234. Concent and emplication by



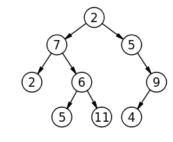




Another example: the term **tree**







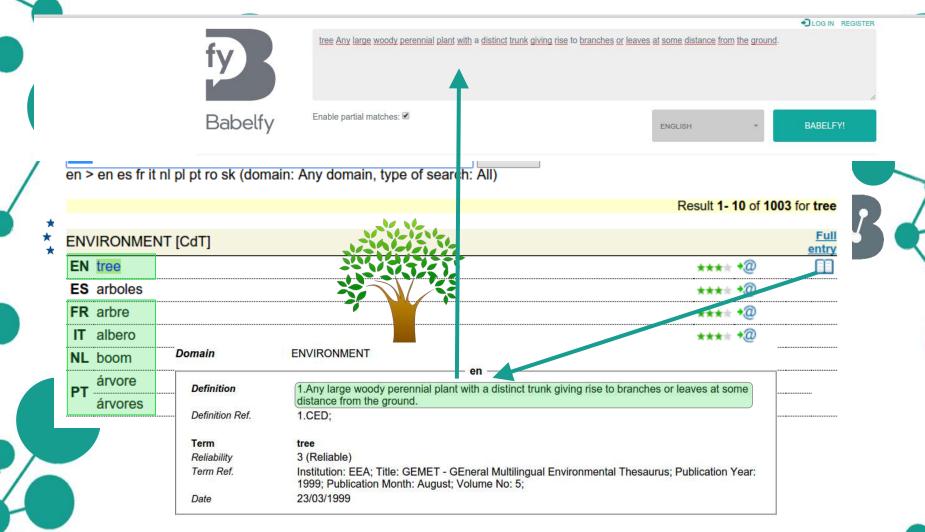


tree Search en > en es fr it nl pl pt ro sk (domain: Any domain, type of search: All) Result 1- 10 of 1003 for tree **Full ENVIRONMENT** [CdT] entry EN tree *** +@ ES arboles FR arbre IT albero **** *****@ **** *@ NL boom árvore Full Electronics and electrical engineering [COM] entry árvores EN tree **** *@ ES árbol **** *@ FR arbre 5 **** *@ If the lemma is the 17 albero *** +@ same, we can still (2) NL boom **** *@ use definitions PT árvore **** *@



en > en es fr it nl pl pt ro sk (domain: Any domain, type of search: All)

Result 1- 10 of 1003 for tree Full ENVIRONMENT [CdT] entry EN tree *** +@ ES arboles FR arbre ***: +@ IT albero **ENVIRONMENT** Domain NL boom árvore Definition 1. Any large woody perennial plant with a distinct trunk giving rise to branches or leaves at some distance from the ground. árvores Definition Ref. 1.CED: Term tree Reliability 3 (Reliable) Term Ref. Institution: EEA; Title: GEMET - GEneral Multilingual Environmental Thesaurus; Publication Year: 1999; Publication Month: August; Volume No: 5; Date 23/03/1999





free Any large woody perennial plant with a distinct trunk giving rise to branches or leaves at some distance from the ground.

Enable partial matches: @

ENGLISH

BABELFY!

to

expanded view | compact view

Any large tree

woody

perennial plant with a

distinct

trunk

giving

rise

branches

tree

A tall perennial woody plant having a main trunk and branches forming a distinct...

woody

Arthur Woody was born in Suches, Georgia.



perennial plant

(botany) a plant lasting for three seasons or more

perennial

Lasting three seasons or more



plant

(botany) a living organism lacking the

distinct

Constituting a separate entity or part



trunk

The main stem of a tree; usually covered with bark: the bole is usually the part that...



rise

A mountain range in the eastern United States extending from Quebec to the Gulf ...



branches

A division of a stem. or secondary stem arising from the main stem of a plant

Legen





free Any large woody perennial plant with a distinct trunk giving rise to branches or leaves at some distance from the ground.

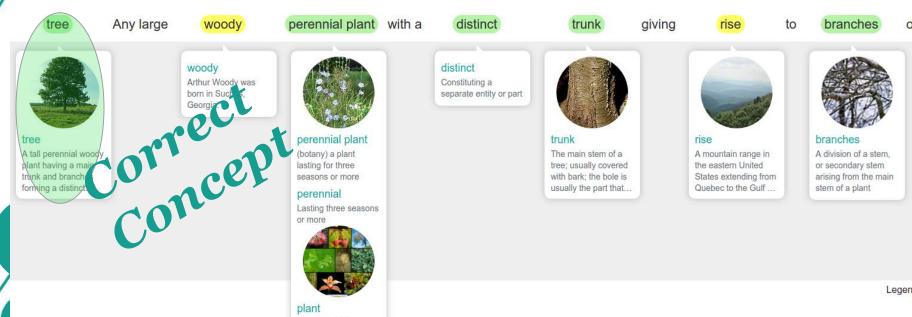
Babelfy

Enable partial matches:

(botany) a living organism lacking the ENGLISH

BABELFY!

expanded view | compact view



Legen





Domain Electronics and electrical engineering

Definition a connected set of branches joining all the nodes of a network without forming a loop Definition Ref.

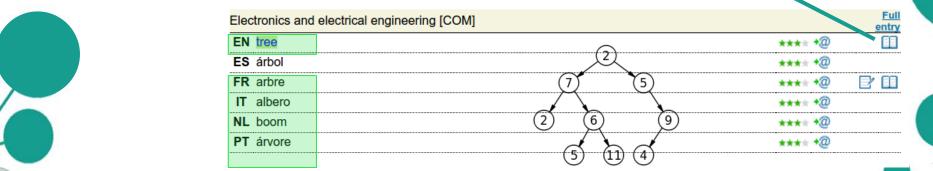
IEV 131-2-9,1978, Electrical and magnetic circuits

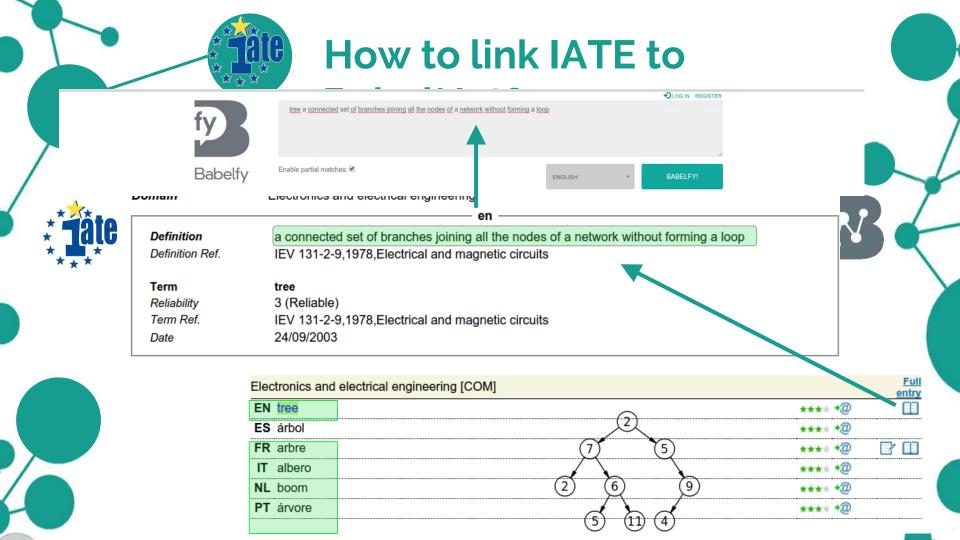
Term tree

3 (Reliable) Reliability

Term Ref. IEV 131-2-9,1978, Electrical and magnetic circuits

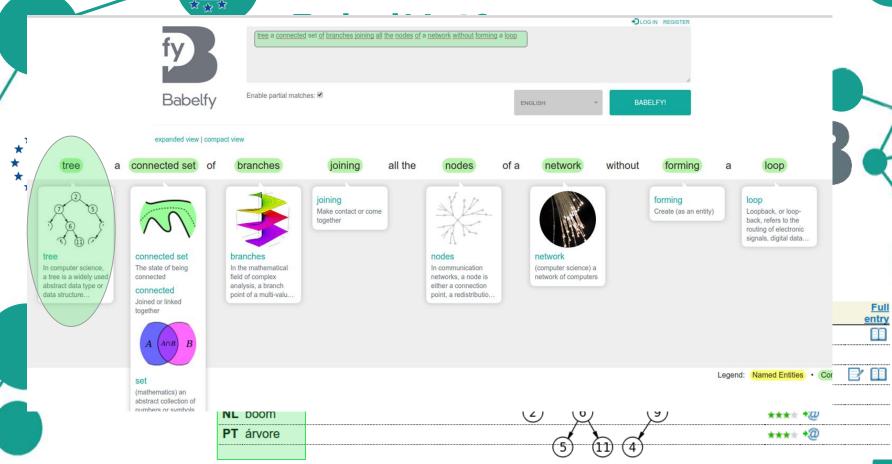
24/09/2003 Date





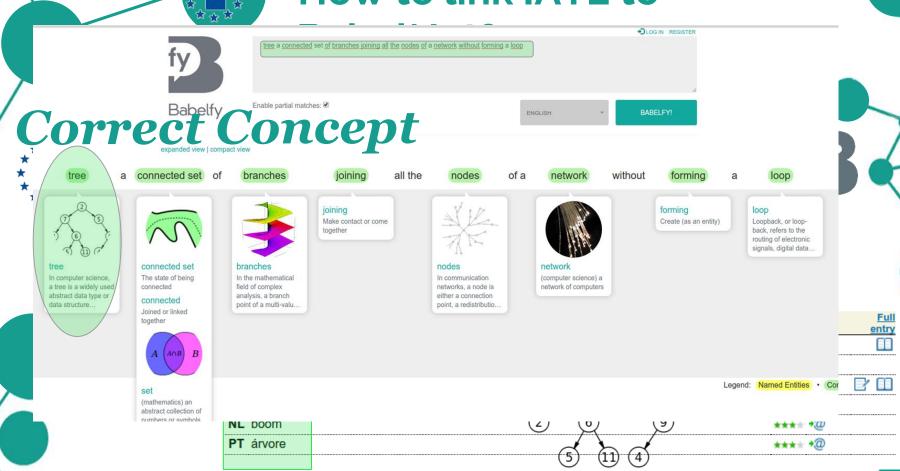


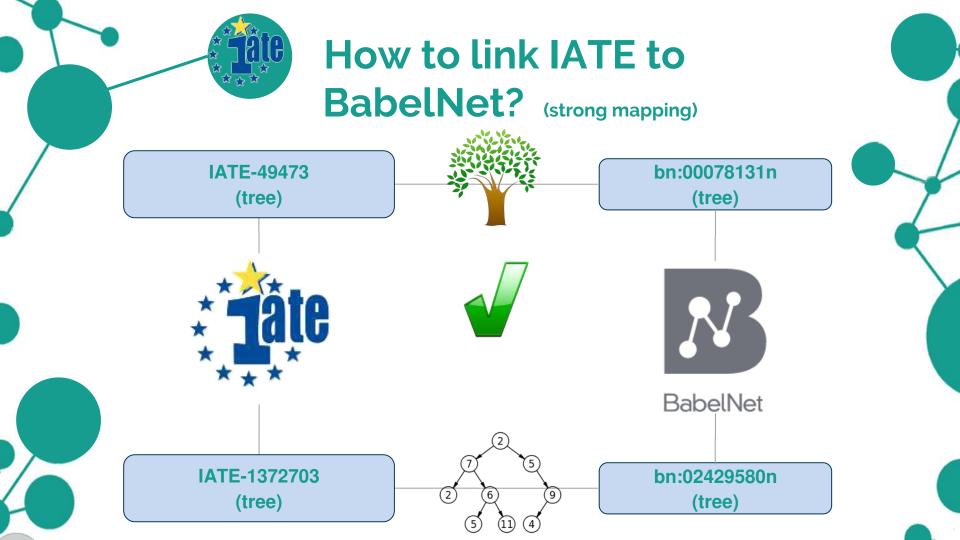
How to link IATE to





How to link IATE to







Conclusion (on IATE)

 The structure of IATE is mostly based on lemmas and textual definitions (even though no definition is available in the dump)



- Strategy: Mapping based on translations and synonyms overlap
- Issue: overlap-based mapping is not enough for concepts with few translations and/or synonyms







Conclusion (on IATE)

- "Stronger" mapping approach IATE-BabelNet:
 - By exploiting **Babelfy** we can potentially improve mapping accuracy and coverage
 - Strategy #1: Disambiguate all synonyms and translations at once using Babelfy with its languageagnostic setting; then select the predominant sense for the lemma of interest;
 - Strategy #2: Disambiguate the target lemma together with its definitions in IATE; a richer context will then enable Babelfy to identify the correct synset associated with the target.





Conclusion (on IATE)

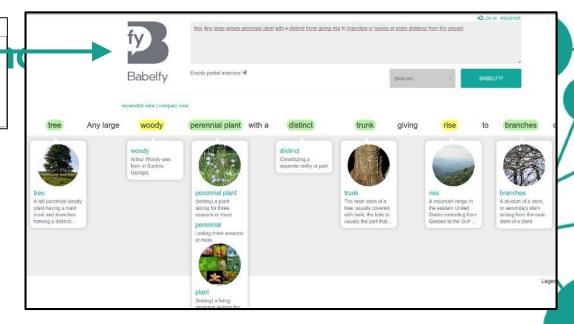
- A third possible mapping approach:
 - Disambiguate both IATE and BabelNet definitions associated with the target lemma; then compare the "bag-of-synsets" associated with each definition and select the IATE concept-BabelNet synset pair whose definition pair maximizes a predefined similarity measure.

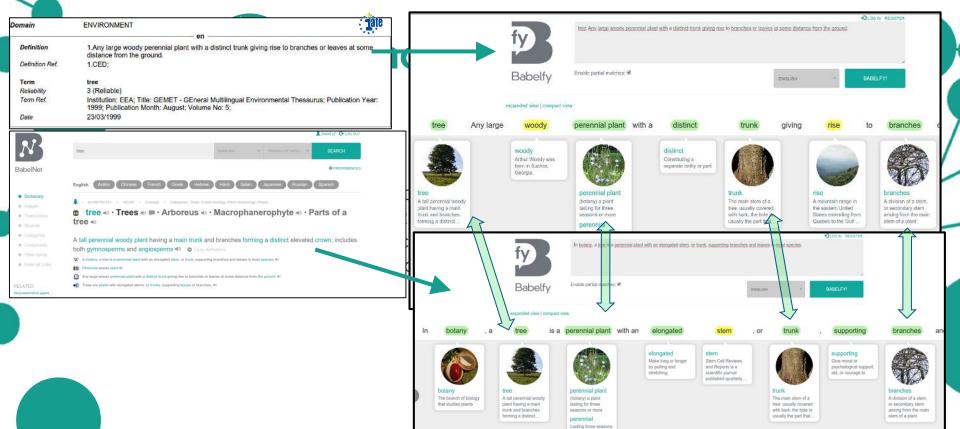




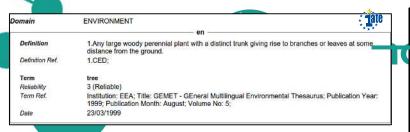
Let us look again at the previous example with the term "tree"

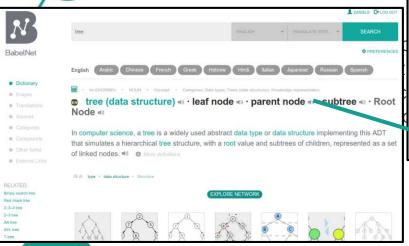
	**
Domain	ENVIRONMENT
15	en
Definition	1.Any large woody perennial plant with a distinct trunk giving rise to branches or leaves at some distance from the ground.
Definition Ref.	1.CED;
Term	tree
Reliability	3 (Reliable)
Term Ref.	Institution: EEA; Title: GEMET - GEneral Multilingual Environmental Thesaurus; Publication Year: 1999; Publication Month: August; Volume No: 5;
Date	23/03/1999

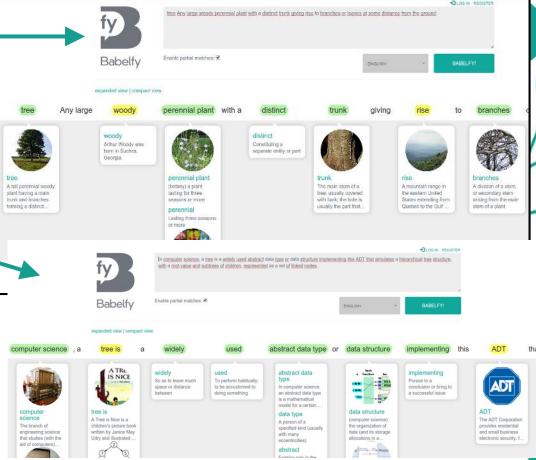




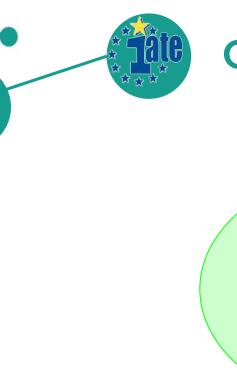
 Higher similarity between definitions = larger number of "shared" senses after disambiguation with Babelfy



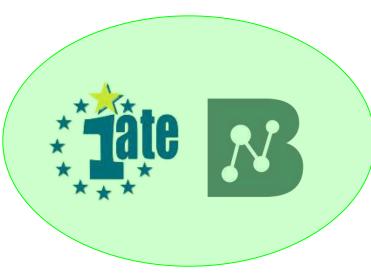




 No shared senses after disambiguation = low (zero) similarity between definitions



Conclusion





Advantages

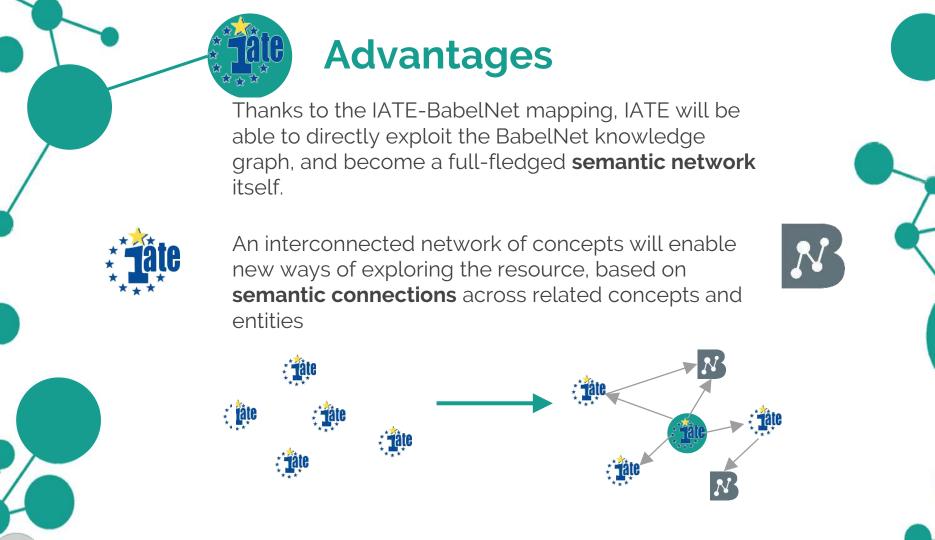
 By integrating with BabelNet, IATE will greatly enlarge its scope and extend its multilingual dictionary from 24 to more than 250 languages.

 At the same time it will enrich BabelNet with additional domain-specific

concepts









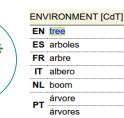
Advantages

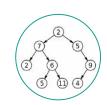
en > en es fr it nl pl pt ro sk (domain: Any domain, type of search: All)

- IATE will benefit from the vast amount of additional information encoded in BabelNet, such as **images**
- It will be possible to identify duplicate and overlapping concepts

Search







EN tree	*** [†] @	88
ES arboles	*** *@	
FR arbre	*** [*] @	
IT albero	**** * @	
NL boom	*** [*] @	
PT árvore	*** [*] @	
árvores	*** <u>*</u>	
Electronics and electrical engineering [COM]		
	***: *@	
EN tree	**** *@	Ent
Electronics and electrical engineering [COM] EN tree ES árbol FR arbre	**** *@ **** *@ **** *@	
EN tree ES árbol	**** *@ **** *@ **** *@	
EN tree ES árbol FR arbre	**** *@ **** *@ **** *@ *** *@	



Result 1- 10 of 1003 for tree











- EuroVoc is a multilingual, multidisciplinary **thesaurus** covering the activities of the EU, the European Parliament in particular
- It contains terms in 23 EU languages (Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish and Swedish), plus Albanian, the language of the FYROM and Serbian
- EuroVoc is managed by the Publications Office, which moved forward to ontology-based thesaurus management and semantic Web technologies following W3C recommendations as well as latest trends in thesaurus standards

Languages	Preferred Terms	Non Preferred Terms (Synonym)	Definitions
BG	6 883	1 681	78
ES	6 883	8 904	229
CS	6 883	14 074	92
DA	6 883	7 594	229
DE	6 883	9 494	227
ET	6 883	7 116	93
EL	6 883	8 092	227
EN	6 883	8 348	233
FR	6 883	8 057	228
ІТ	6 883	10 592	228



- EuroVoc is already a semantic network
- EuroVoc is based on domains (21 overall)
- Each domain is further subdivided into sub-domains (127 overall)



- EuroVoc is already a semantic network;
- EuroVoc is based on domains (21 overall)
- Each domain is further subdivided into sub-domains (127 overall)
 - POLITICS
 - INTERNATIONAL RELATIONS
 - EUROPEAN UNION
 - LAW
 - SCIENCE
 - ...

- political framework
- political party
- electoral procedure and voting
- parliament
- parliamentary proceedings
- .



- A EuroVoc entry contains:
 - synonyms
 - translations
 - definitions in multiple languages
 - domains
 - sub-domains
 - relations

ethnography

SN

The branch of anthropology that deals with the scientific description of individual human societies.

36 SCIENCE

MT 3611 humanities

BT1 ethnology

BT2 social sciences
RT ethnic group [2821]

URI http://eurovoc.europa.eu/6394

Has Exact Match

Ethnography (ECLAS)

LANGUAGE EQUIVALENTS

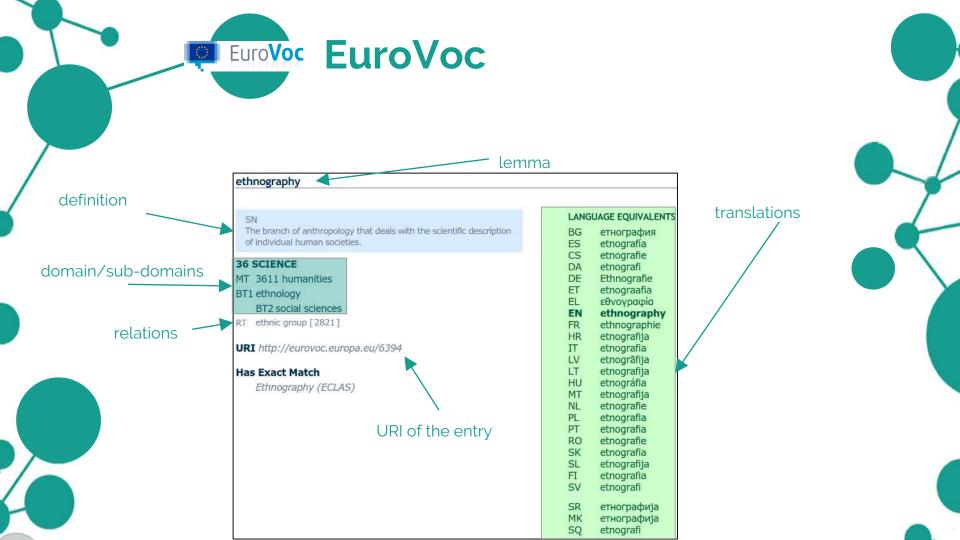
BG етнография ES etnografia CS etnografie DA etnografi DE Ethnografi ET etnografia EL εθνογραφία

EN ethnography
FR ethnographie
HR etnografija
IT etnografia

LV etnogrāfija LT etnogrāfija HU etnogrāfija NL etnografija NL etnografija PL etnografia PT etnografia RO etnografia SK etnografia

SL etnografija FI etnografia SV etnografi

SR етнографија MK етнографија SQ etnografi





EUTOVOC Linking EuroVoc to BabelNet

 Goal: To automatically (and semantically) link EuroVoc to BabelNet using a language- and resource-agnostic approach





BabelNet



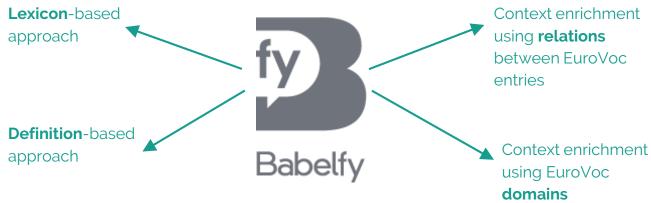


EuroVoc How to link EuroVoc to BabelNet? (strong mapping)



Again we can leverage **Babelfy**, our joint graphbased approach to multilingual Entity Linking and **Word Sense Disambiguation.**





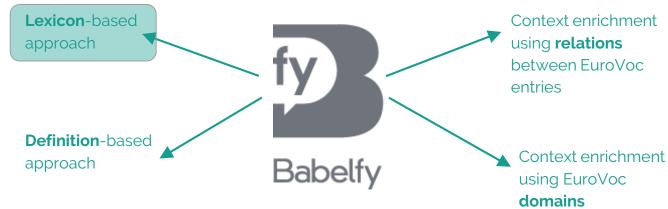


EuroVoc How to link EuroVoc to BabelNet? (strong mapping)



Again we can leverage **Babelfy**, our joint graphbased approach to multilingual Entity Linking and **Word Sense Disambiguation.**







How to link EuroVoc to BabelNet? (strong mapping - 1)



Lexicon-based approach









EuroVoc How to link EuroVoc to BabelNet? (strong mapping - 1)



Lexicon-based approach

Target word: cement



cement cement industry

68 INDUSTRY

MT 6831 building and public works

BT1 building materials BT2 building industry

URI http://eurovoc.europa.eu/5993

Has Exact Match

Cement (AGROVOC)

Cement (ECLAS)

cement (GEMET) cement (UNBIS)

Has Related Match

Cement industry (ECLAS) cement industry (GEMET)

cement manufacture (GEMET)

cement industry (UNBIS)

LANGUAGE EQUIVALENTS

BG цимент cemento cement

DA cement DF Zement tsement

τσιμέντο EL EN cement

ciment cement IT cemento

cements LT cementas HU cement

MT siment cement cement

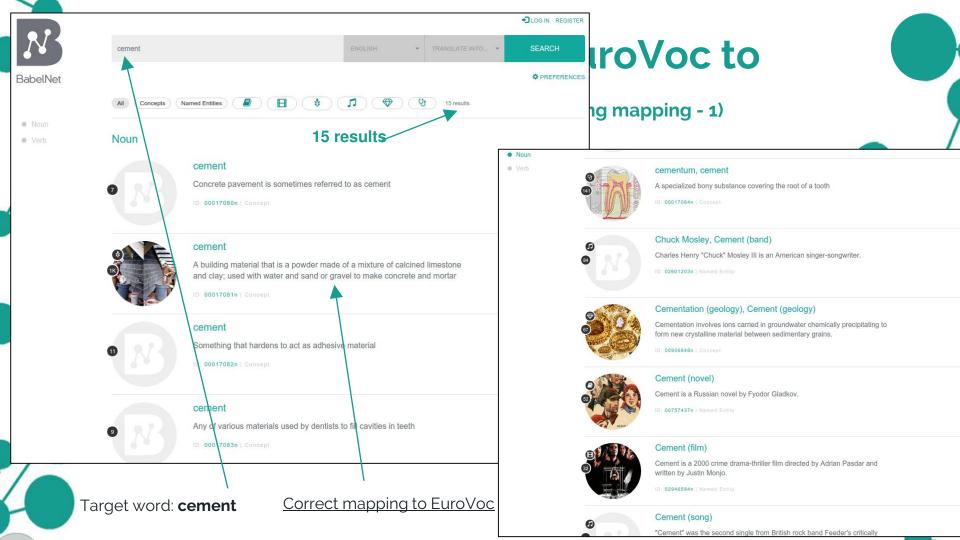
cimento ciment SK cement

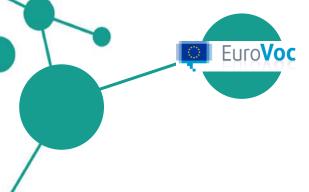
cement sementti

cement

цемент

цемент cimento





How to link EuroVoc to BabelNet? (strong mapping - 1)









How to link EuroVoc to BabelNet? (strong mapping - 1)



UF cement industry

68 INDUSTRY

MT 6831 building and public works

BT1 building materials

BT2 building industry

URI http://eurovoc.europa.eu/5993

Has Exact Match

Cement (AGROVOC)

Cement (ECLAS)

cement (GEMET)

cement (UNBIS)

Has Related Match

Cement industry (ECLAS)

cement industry (GEMET)

cement manufacture (GEMET)

cement industry (UNBIS)

LANGUAGE EQUIVALENTS

BG цимент

ES cemento

CS cement

DA cement

DE Zement

ET tsement EL τσιμέντο

EN cement

FR ciment HR cement

IT cement

LV cements

LT cementas

MT siment

NL cement

PL cement PT cimento

RO ciment SK cement

SK cement SL cement

FI sementti SV cement

SR цемент

MK цемент SQ cimento fy

Babelfy





MIND

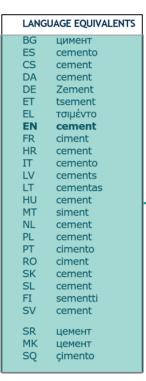
THE LEXICAL

EUroVoc Building block of the semantic web



EuroVoc How to link EuroVoc to BabelNet? (strong mapping - 1)

MIND THE LEXICAL EUroVoc Building block of the semantic web



Lexicon-based approach - cement





Babelly is an output of the MultiUEDI ERC Starting Grant No. 208204. Concept and application by Arriban Manuard Robarto Navago Bahally and the API are immediately a Common Common Attribution-Non Commercial-Share Alike 3.0 License. For any commercial use, please contact us.









EuroVoc How to link EuroVoc to BabelNet? (strong mapping - 1)

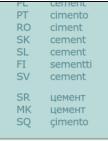


цимент, cement, cement, cement, zement, zement, tsement, roiµévro, cement, ce cement, cimento, ciment, cement, sementti, cement, цемент, цемент, cimento

Enable partial matches:

AGNOSTIC

BABELEY!













expanded view | compact view

Zement τσιμέντο ciment cemento cement tsement cement цимент cement



A building material that is a powder made of a mixture of calcined limestone...



A building material that is a powder made of a mixture of calcined limestone...



A building material that is a powder made of a mixture of calcined limestone...



A building material that is a powder made of a mixture of calcined limestone...



Zement A building material that is a powder made of a mixture of calcined limestone...



AGNOSTIC

tsement A building material that is a powder made of a mixture of calcined limestone...



A building material that is a powder made of a mixture of calcined limestone...



cement A building material that is a powder made of a mixture of calcined limestone...



ciment A building material that is a powder made of a mixture of calcined limestone...







expanded view | compact view

cement



calcined limestone



cement

A building material that is a powder made of a mixture of calcined limestone



sementti

A building material that is a powder made of a mixture of calcined limestone



cement

cement A building material that is a powder made of a mixture of calcined limestone



цемент

A building material that is a powder made of a mixture of calcined limestone



цемент

цемент A building material that is a powder made of a mixture of calcined limestone



cimento

A building material that is a powder made of a mixture of calcined limestone



How to link EuroVoc to BabelNet? (strong mapping - 1)

EuroVoc-5993 (cement)





bn:00017081n (cement)



BabelNet

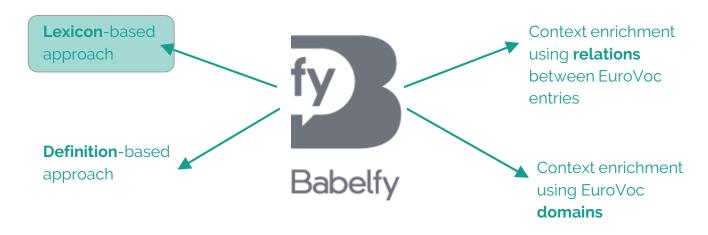


How to link EuroVoc to BabelNet? (strong mapping)



We can leverage **Babelfy**, our joint graph-based approach to multilingual **Entity Linking** and **Word Sense Disambiguation**.





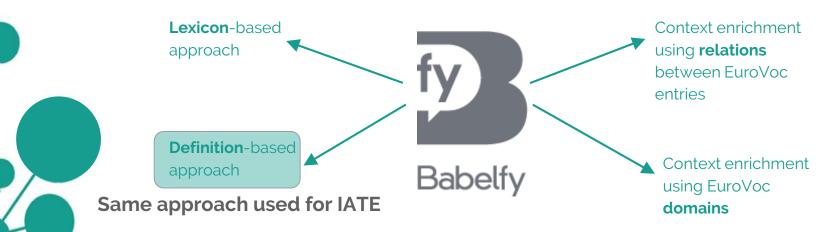


How to link EuroVoc to BabelNet? (strong mapping)



We can leverage **Babelfy**, our joint graph-based approach to multilingual **Entity Linking** and **Word Sense Disambiguation**.





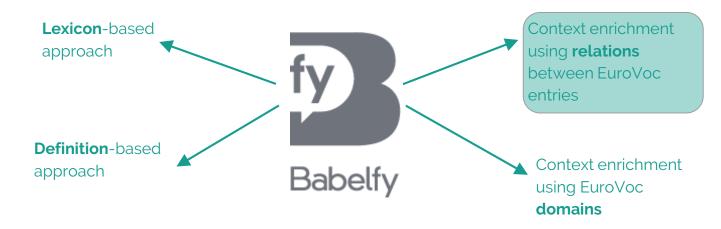


How to link EuroVoc to BabelNet? (strong mapping)



We can leverage **Babelfy**, our joint graph-based approach to multilingual **Entity Linking** and **Word Sense Disambiguation**.







How to link EuroVoc to BabelNet? (strong mapping - 3)



Context enrichment using **relations** between EuroVoc entries

Target word: **missile**







EuroVoc How to link EuroVoc to

missile





Context enrichment using **relations** between EuroVoc entries

Target word: missile



UF nuclear missile

08 INTERNATIONAL RELATIONS

MT 0821 defence

BT1 military equipment

NT1 anti-missile defence

NT1 ballistic missile

NT1 guided missile

NT1 intercontinental missile

RT Euro-missile [0816] nuclear technology [6621]

URI http://eurovoc.europa.eu/1944

Has Exact Match

Missiles (ECLAS)

LANGUAGE EQUIVALENTS

BG ракета ES misil CS

raketa DA missil

Rakete DE

ΕT rakett πύραυλοι

EN missile

missile HR projektil

IT missile

raķete raketa LT

rakéta HU MT missila

raket NL

broń rakietowa PL

míssil rachetă RO

SK raketa

izstrelek FΙ ohjus

SV robot

SR ракета MK проектил

SQ raketë



Target word: **missile**



Correct mapping to EuroVoc



Noun 5 results



missile

A rocket carrying a warhead of conventional or nuclear explosives; may be ballistic or directed by remote control

ID: 00055322n | Concept



projectile, missile

A weapon that is forcibly thrown or projected at a targets but is not selfpropelled

ID: 00055323n | Concept



Missile (Ghost Trick)

Missile is a fictional character and a major character of the Nintendo DS video game Ghost Trick: Phantom Detective.

ID: 02432795n | Named Entity



Missile (song)

"Missile" is a song by IAMX from the album Kiss + Swallow.

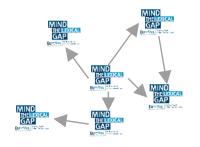
ID: 01703991n | Named Entity

Missile (film)



How to link EuroVoc to BabelNet? (strong mapping - 3)

















EuroVoc How to link EuroVoc to BabelNet? (strong mapping - 3)





missile

UF nuclear missile

08 INTERNATIONAL RELATIONS

MT 0821 defence

BT1 military equipment

NT1 anti-missile defence

NT1 ballistic missile NT1 quided missile

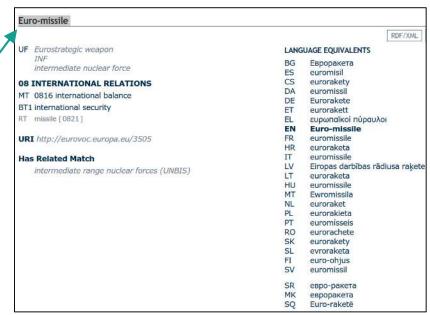
NT1 intercontinental missile

RT Euro-missile [0816] nuclear technology [6621]

URI http://eurovoc.europa.eu/1944

Has Exact Match

Missiles (ECLAS)







EuroVoc How to link EuroVoc to BabelNet? (strong mapping - 3)





missile

UF	nuclear missile
08	INTERNATIONAL RELATIONS
MT	0821 defence
BT1	military equipment
	anti-missile defence
–	ballistic missile
	guided missile intercontinental missile
	Euro-missile [0816]
KI	nuclear technology [6621]
	nuclear technology [0021]

URI http://eurovoc.europa.eu/1944

Has Exact Match Missiles (ECLAS)

Euro-missile		
UF Eurostrategic weapon INF Intermediate nuclear force 08 INTERNATIONAL RELATION MT 0816 international balance BT1 international security RT missil [0821] URA http://eurovoc.europa.eu/3 Has Related Match Intermediate range nuclear	MT 6621 electrical and nuclear industries BT1 nuclear industry NT1 fuel enrichment NT1 fuel reprocessing NT1 nuclear chemistry NT1 nuclear fission NT1 nuclear fusion	RDF/XM LANGUAGE EQUIVALENTS BG ЯДРЕНА ТЕХНОЛОГИЯ ES tecnologia nuclear CS jaderná technologie DA atomteknologi DE Kerntechnologie ET tuumatehnoloogia EL пирлукή техуолоуіа EN nuclear technology FR technologie nucléaire HR nuklearna tehnologija IT tecnologia nucleare LV kodoltehnologija LT branduolinė technologija HU nukleáris technologia MT teknoloģija nukleari NL nucleaire technologie PL technika jadrowa PT tecnologia nuclear RO tehnologie nucleară
	therapeutics [2841] URI http://eurovoc.europa.eu/4422	SK jadrová technológia SL jedrska tehnologija FI ydinteknologia SV kärnteknologi
	Has Exact Match nuclear technology (UNBIS) Has Related Match	SR нуклеарна технологија МК нуклеарна технологија SQ teknologji bërthamore

Particle accelerators (ECLAS)

EuroVoc Graph



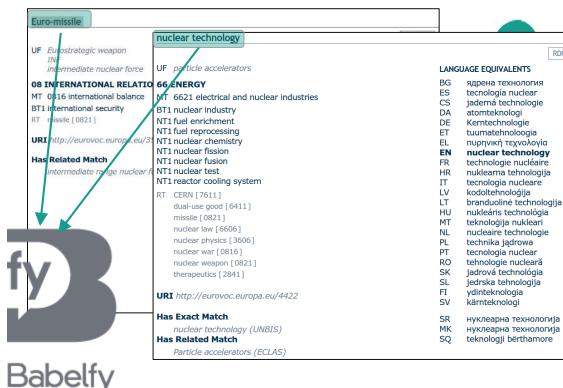
EuroVoc How to link EuroVoc to

BabelNet? (strong mapping - 3)

RDF/XML





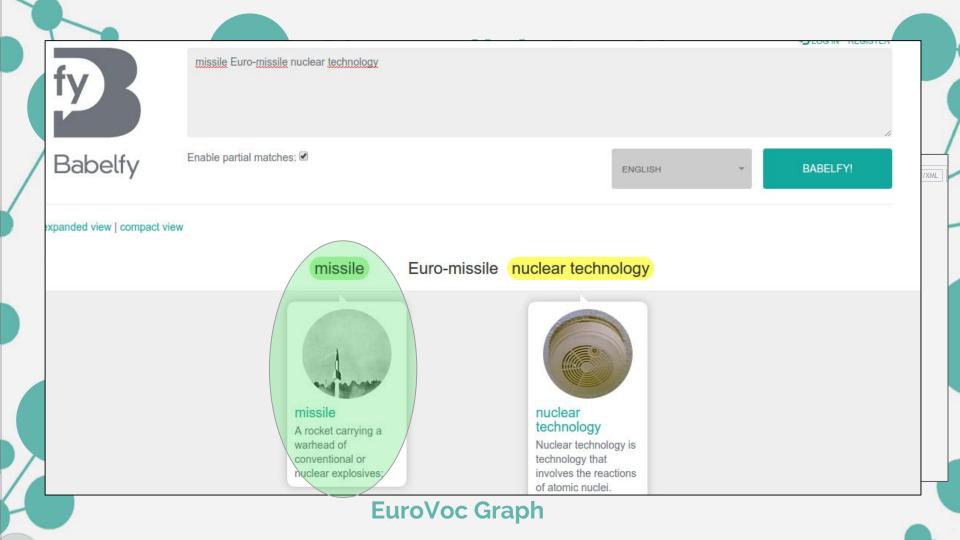


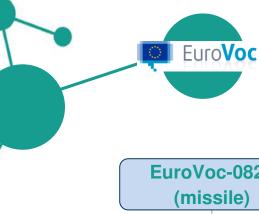


EuroVoc How to link EuroVoc to

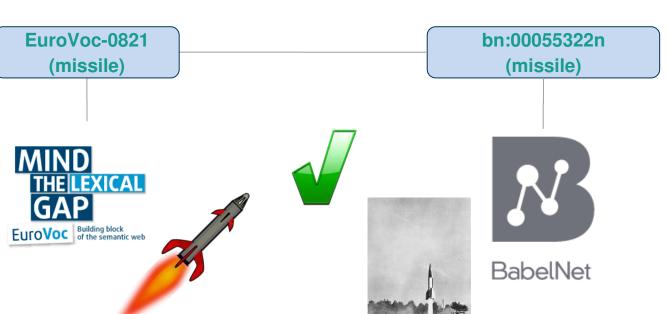
BabelNet? (strong mapping - 3)







EuroVoc How to link EuroVoc to BabelNet? (strong mapping - 3)



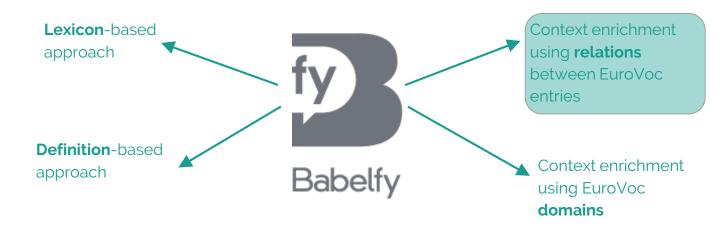


How to link EuroVoc to BabelNet? (strong mapping)



We can leverage **Babelfy**, our joint graph-based approach to multilingual **Entity Linking** and **Word Sense Disambiguation**.





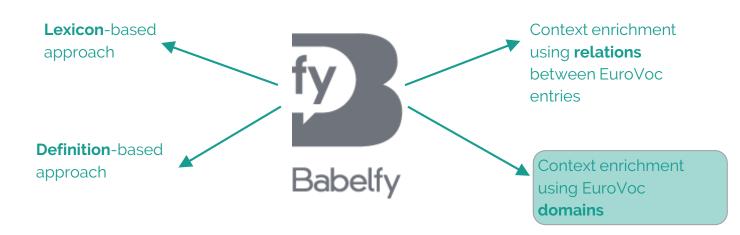


How to link EuroVoc to BabelNet? (strong mapping)



We can leverage **Babelfy**, our joint graph-based approach to multilingual **Entity Linking** and **Word Sense Disambiguation**.



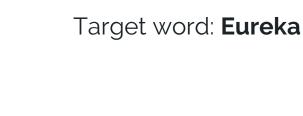




How to link EuroVoc to BabelNet? (strong mapping - 4)



Context enrichment using FuroVoc **domains**









EuroVoc How to link EuroVoc to

Ba

Eureka



Context enrichment using EuroVoc domains

Target word: **Eureka**



UF Eureka charter Eureka ministerial conference Eureka programme Eureka project European Research Coordination Agency

64 PRODUCTION, TECHNOLOGY AND RESEARCH

MT 6416 research and intellectual property BT1 research policy

URI http://eurovoc.europa.eu/3654

LANGUAGE EQUIVALENTS

BG Eypeka
ES Eureka
CS program Eureka
DA Eureka

DE Eureka ET Eureka

EL Eureka **EN Eureka**

FR Eureka HR Eureka IT Eureka

LV Eureka LT Eureka HU Eureka

MT Eureka NL Eureka PL Eureka

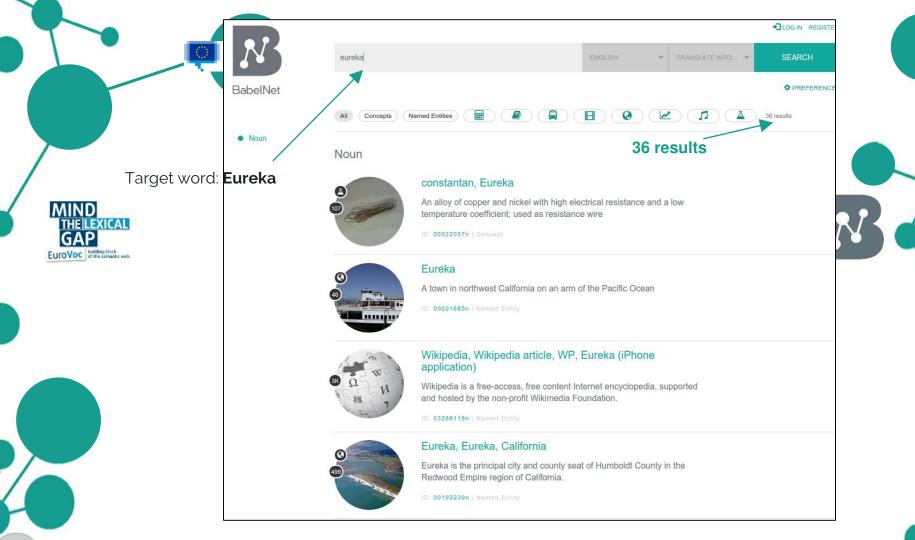
T Eureka O Eureka

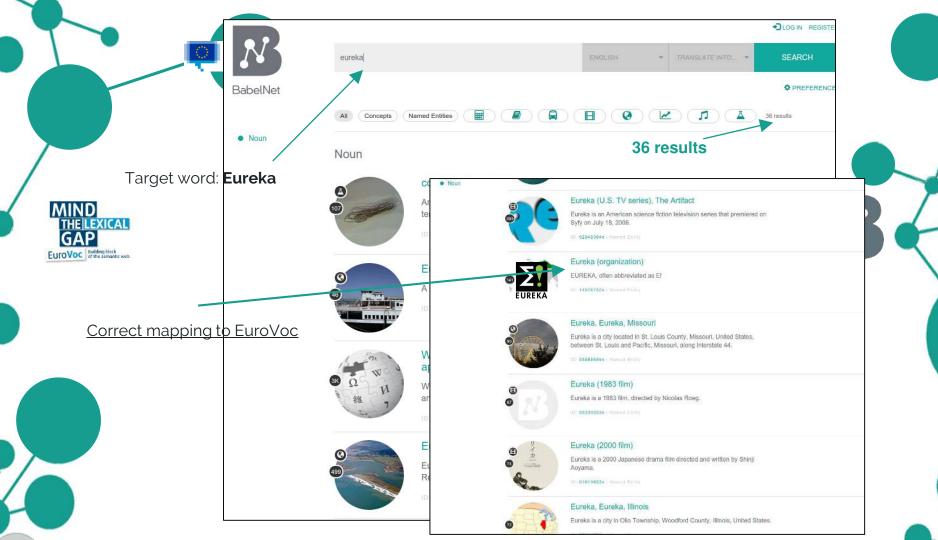
SK Eureka SL Eureka T Fureka

SV Eureka

SR Eypeка МК Eypeка

SQ Eureka







How to link EuroVoc to BabelNet? (strong mapping - 4)



Disambiguating Eureka with no context:

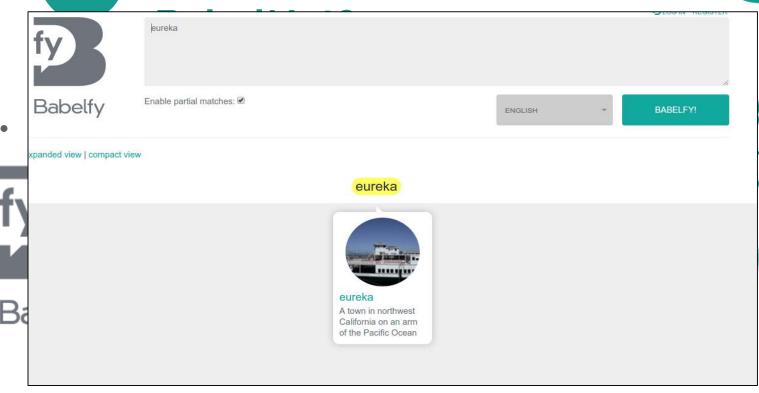


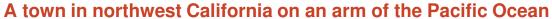
Babelfy





EuroVoc How to link EuroVoc to



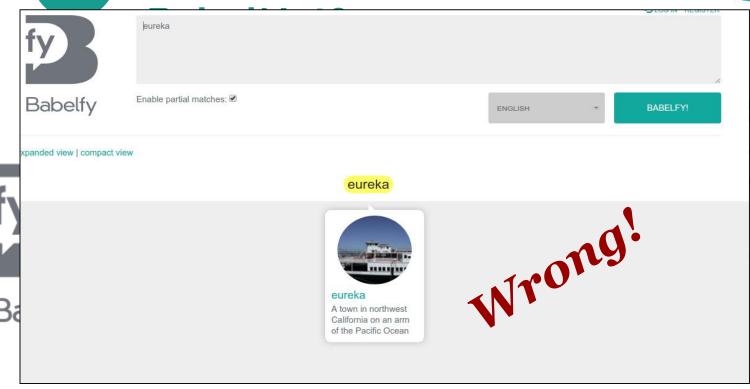








EuroVoc How to link EuroVoc to



A town in northwest California on an arm of the Pacific Ocean







How to link EuroVoc to BabelNet? (strong mapping - 4)





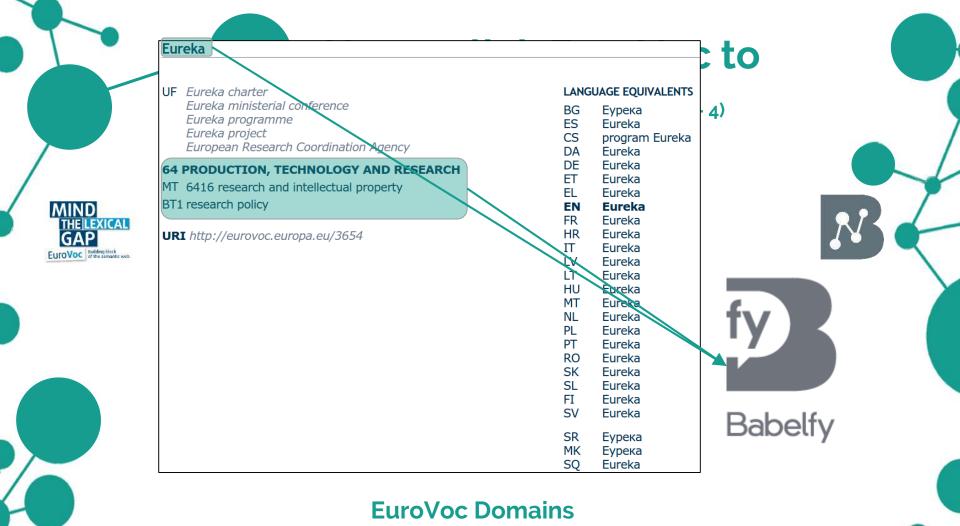


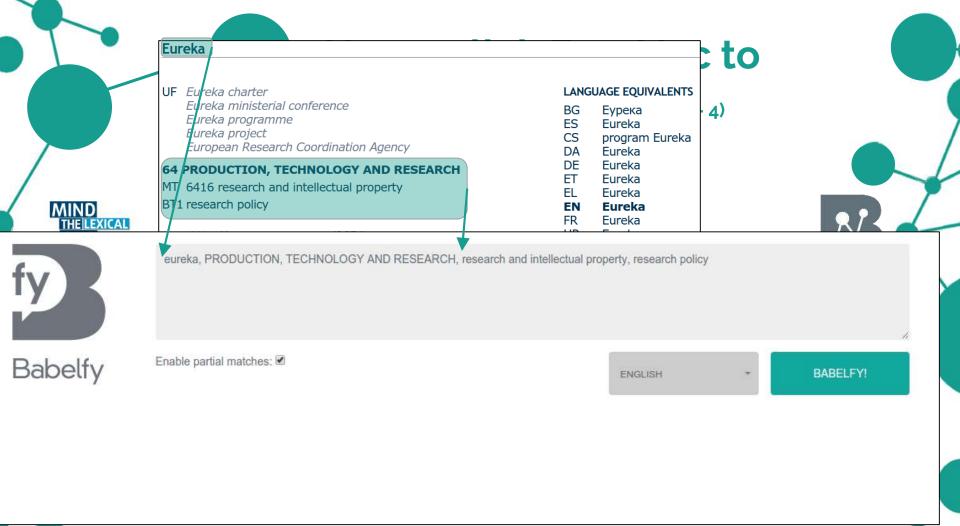














DLOG IN REGISTER



Enable partial matches:

ENGLISH

BABELFY!

expanded view | compact view

eureka

PRODUCTION , TECHNOLOGY AND RESEARCH ,

research

and intellectual property , research policy



eureka

EUREKA, often abbreviated as E! PRODUCTION

(economics) manufacturing or mining or growing something (usually i... Agency for Science. Tec and Resea

TECHNOLOGY AND RESEARCH

The Agency for Science, Technology and Research is a statutory board und...



TECHNOLOGY

An engineering



research

Systematic investigation to establish facts



intellectual property

Intangible property that is the result of creativity (such as patents or trademar...

intellectual

Of or associated with or requiring the use of the mind



ity of Suss

re Policy

research policy

Science and Technology Policy Research is an academic center at...



research

Systematic investigation to establish facts

Name



EuroVoc How to link EuroVoc to

BabelNet? (strong mapping - 4)





bn:14976752n (Eureka)







BabelNet

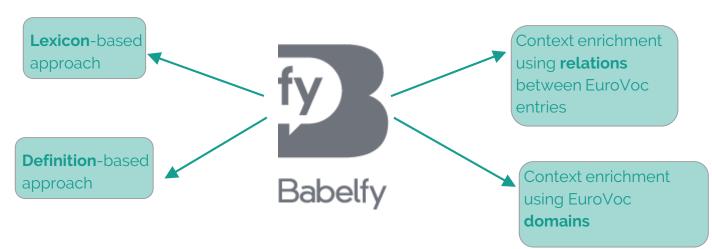


How to link EuroVoc to BabelNet? (strong mapping)



A more advanced mapping approach could leverage **all these basic strategies** at the same time to improve accuracy







Euro Voc Conclusion (on Euro Voc)

- Various basic approaches for the BabelNet-EuroVoc mapping:
 - Lexicon-based approach
 - Exploits lemmas and translations of a EuroVoc entry
 - Definition-based approach
 - Exploits the textual definition as useful context for disambiguation
 - Context enrichment using relations
 - Exploits relations between EuroVoc entries as context
 - Context enrichment using domains
 - Exploits EuroVoc domains and subdomains
- More advanced techniques: semantic similarities based on vector-space models, etc.



By integrating with BabelNet, EuroVoc (like IATE) will greatly enlarge its scope and extend its multilingual dictionary to more than **250** languages.









- EuroVoc is already a semantic network, but its scope is relatively limited: by integrating with BabelNet, many more connections will be introduced, enlarging EuroVoc's scope and dimension
- New content (e.g. images) associated with EuroVoc entries





missile
A rocket ca
ballistic or o

ID: 00055322

2 relations

949 semantic relations

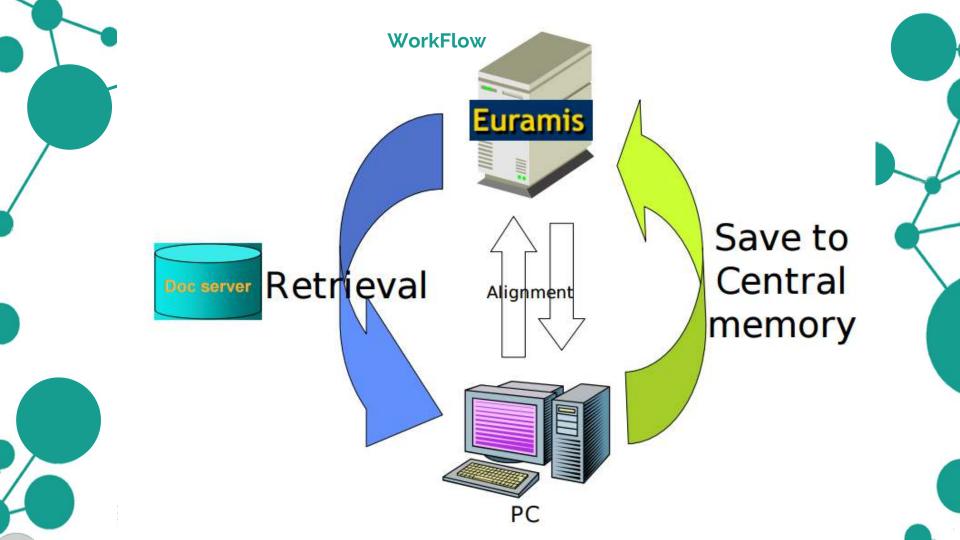


Euramis

European Advanced Multilingual Information System



- Euramis is a system of **databases** or **memories** (and associated tools) fed by the various institutions jointly (Legis-Process) or individually.
- It acts as a multilingual, multidirectional repository of clearly labelled **equivalent phrases** ("segments") belonging to official EU documents allowing their re-use in translation in all European Institutions.
- The system also contains a set of **interfaces** to manage access rights for sentence and document management, based on the user's login identification, for both Commission staff and other institution's staff.





Alignment

ource sentence: English	Translation sentence: Italian	6/516
THE COMMISSION OF THE EUROPEAN COMMUNITIES,	LA COMMISSIONE DELLE COMUNITÀ EUROPEE,	< >>
Having regard to the Treaty establishing the European Community,	visto il trattato che istituisce la Comunità europea,	Merge
Having regard to Regulation No 79/65/EEC of the Council of 15 June 1965 setting up a network for the collection of accountancy data on the incomes and business operation of agricultural holdings in the European Economic Community, and in particular Article 7(3) thereof,	visto il regolamento n. 79/65/CEE del Consiglio, del 15 giugno 1965, relativo all'istituzione di una rete d'informazione contabile agricola sui redditi e sull'economia delle aziende agricole nella CEE, in particolare l'articolo 7, paragrafo 3,	Split Delete
Whereas:	considerando quando segue:	Undo Redo
Commission Regulation (EEC) No 2237/77 lays down the contents of the farm return to be used.	Il regolamento (CEE) n. 2237/77 della Commissione stabilisce il contenuto della scheda aziendale da utilizzare.	Close



- **ENGLISH:** These imports competed with each other and with the like product manufactured by the Community industry.
- **SPANISH:** Estas importaciones compitieron entre sí y con el producto similar manufacturado por la industria de la Comunidad.

• **ENGLISH:** The Commission, therefore, has excluded this company from the definition of the 'the Community industry` for the purpose of its injury determination.

• **SPANISH:** La Comisión, por lo tanto, ha excluido a esta empresa de la definición de la industria de la Comunidad a efectos de la determinación del perjuicio.

 Keyword search is currently used to browse the Euramis content:





These **imports** competed with each other and with the like product manufactured by the Community industry.

The volume of dumped **imports** into the Community of the product concerned originating...

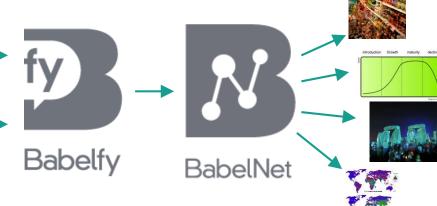
Euramis Semantic search with Euramis

• Keyword search is currently used to browse the Euramis content:



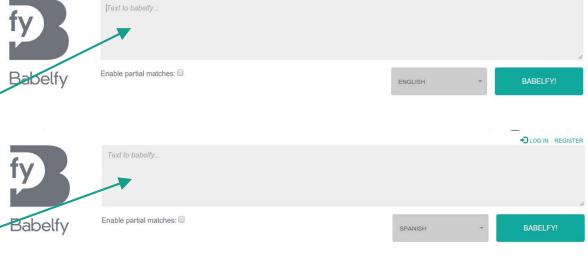
We can use BabelNet synsets in place of plain keywords!

- What can we do with Euramis?
 - Being a repository of parallel text, Euramis does **not** associate each sentence with a particular lemma = no strict mapping Euramis-BabelNet is possible at the sentence level
 - But: we can still exploit **Babelfy** to **index** Euramis sentences with the sense inventory of **BabelNet**
- ENGLISH: These imports competed with each other and with the like product manufactured by the Community industry.
- SPANISH: Estas importaciones compitieron entre sí y con el producto similar manufacturado por la industria de la Comunidad.



3LOG IN REGISTER

- ENGLISH: These imports competed with each other and with the like product manufactured by the Community industry.
- SPANISH: Estas importaciones compitieron entre sí y con el producto similar manufacturado por la industria de la Comunidad.





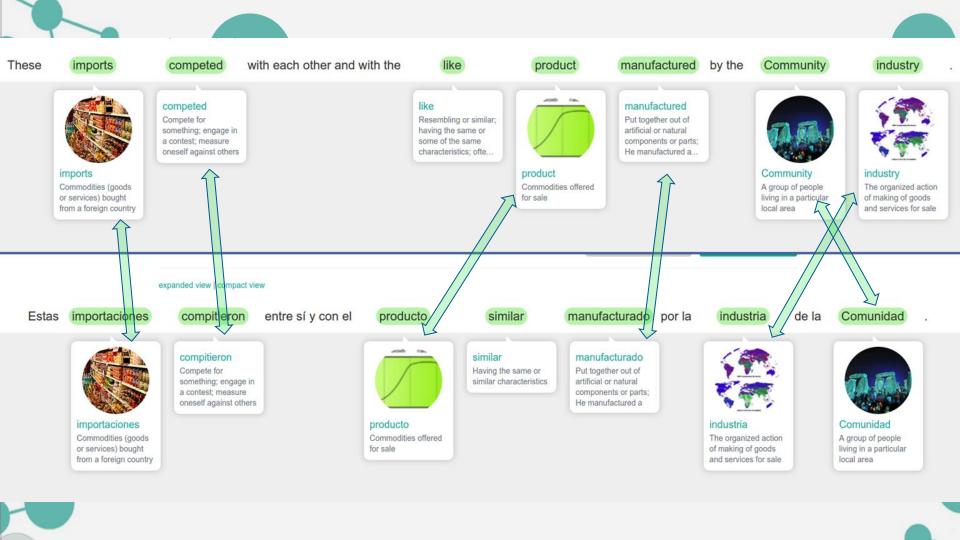
ABOUT
PUBLICATIONS
DOWNLOADS
API GUIDE

Babelty is an output of the MultiJEDI ERC Starting Grant No. 299234. Concept and application by Andrea Moro and Roberto Navigil. Babelty and its API are licensed under a Creative Commons Attribution-Non Commercial-Share Alike 3.0 License. For any commercial use, please contact us.









- **ENGLISH:** These <u>imports</u> <u>competed</u> with each other and with the like <u>product manufactured</u> by the <u>Community industry</u>.
- **SPANISH:** Estas <u>importaciones</u> <u>compitieron</u> entre sí y con el <u>producto</u> similar <u>manufacturado</u> por la <u>industria</u> de la <u>Comunidad</u>.



importaciones

Commodities (goods or services) bought from a foreign country

compitieron

Compete for something; engage in a contest; measure oneself against others



producto

Commodities offered for sale

manufacturado

Put together out of artificial or natural components or parts; He manufactured a



industria

The organized action of making of goods and services for sale



Comunidad

A group of people living in a particular local area

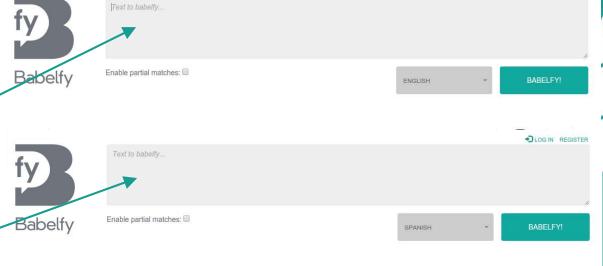


• **ENGLISH:** The Commission, therefore, has excluded this company from the definition of the 'the Community industry` for the purpose of its injury determination.

 SPANISH: La Comisión, por lo tanto, ha excluido a esta empresa de la definición de la industria de la Comunidad a efectos de la determinación del perjuicio.

DLOG IN REGISTER

- ENGLISH: The Commission, therefore, has excluded this company from the definition of the 'the Community industry' for the purpose of its injury determination.
- SPANISH: La Comisión, por lo tanto, ha excluido a esta empresa de la definición de la industria de la Comunidad a efectos de la determinación del perjuicio.



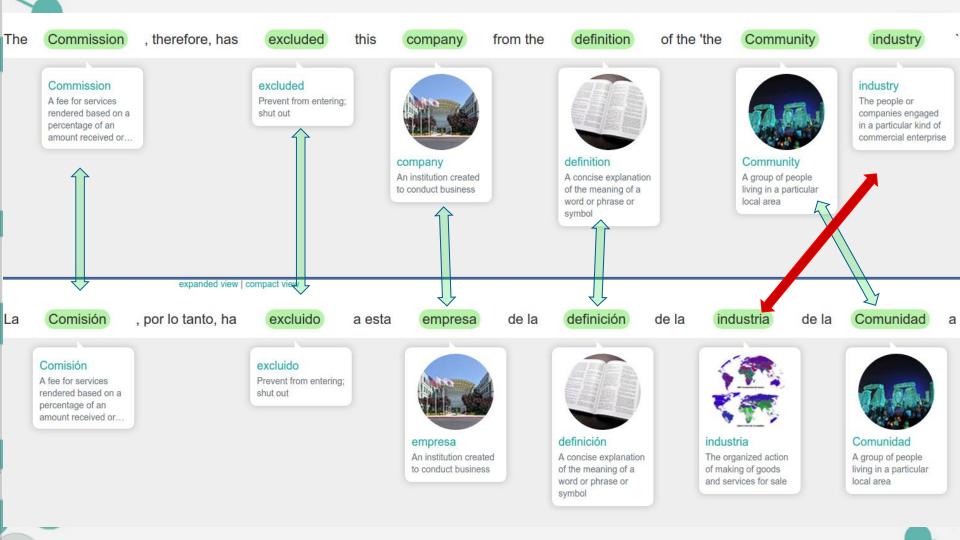


Babetly is an output of the MultiJEDL ERC Starting Grant No. 259234. Concept and application by Andrea Moro and Roberto Navigli. Babetly and its API are licensed under a Creative Commons Attribution-Non Commercial-Share Alike 3.0 License. For any commercial use, please contact us:

(1) (2) (2)





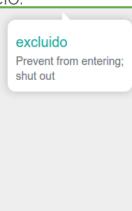


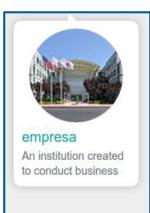
• **ENGLISH:** The <u>Commission</u>, therefore, has <u>excluded</u> this <u>company</u> from the <u>definition</u> of the 'the <u>Community</u> industry` for the purpose of its injury determination.

SPANISH: La <u>Comisión</u>, por lo tanto, ha <u>excluido</u> a esta <u>empresa</u> de la <u>definición</u> de la industria de la <u>Comunidad</u> a efectos de la determinación <u>del perjuicio</u>.

Comisión

A fee for services rendered based on a percentage of an amount received or...











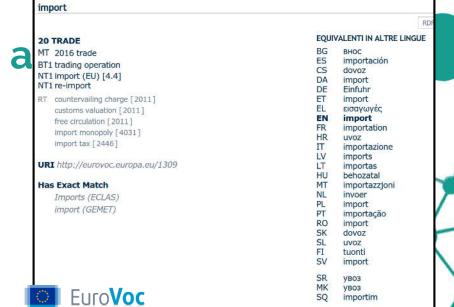
- **Euramis** is neither a glossary nor a thesaurus: as such, a strict mapping from Euramis entries to BabelNet synsets is inherently impossible
- Instead of mapping, we can however **disambiguate** each sentence
- Thanks to the Euramis alignments and the multilingual nature of BabelNet, we can then identify the most frequent Euramis concepts across languages;
- We can **index** each sentence with the corresponding concepts identified through disambiguation, enabling concept-based browsing and search of the Euramis resource.



 As a result of all mapping and indexing procedures we just saw, the three examined resources (IATE, EuroVoc and Euramis) can eventually connected to each other through BabelNet!



Definition	en	
	goods or services purchased from another country "imports". A Dictionary of Business and Management (5 ed.). Edited by Jonathan Law. Oxford University Press. Current Online Version: 2014, www.oxfordreference.c [2.10.2014]	
Definition Ref.		
Term	imports	
Reliability	3 (Reliable)	
Term Ref.	Implementing Regulation (EU) No 1039/2014 establishing the standard import values for determining the entry price of certain fruit and vegetables, eur-lex.europa.eu/leg	
Date	02/10/2014	
Term	importations	
Reliability	2 (Minimum reliability)	
Term Ref.	Implementing Regulation (EU) No 1001/2013 amending Annex I to Regulation (EEC) No 2658/87 on the tariff and statistical nomenclature and on the Common Customs Tariff, 32013R1001:EN:NOT* >32013R1001/EN:D	
Date	02/10/2014	
	it	
Definition	merci, servizi o capitali provenienti dall'estero e immessi nel territorio di uno Stato	
Definition Ref.	COM-IT sulla base di: Enciclopedia Treccani.it, Importazione, www.treccani.it/encic [31.10.2014]	
Term	importazioni	



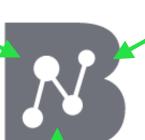


- These imports competed with each other and with the like product manufactured by the Community industry.
- Estas importaciones compitieron entre sí y con el producto similar manufacturado por la industria de la Comunidad.





PINET as Se MT 2016 trade BT1 trading (FI)



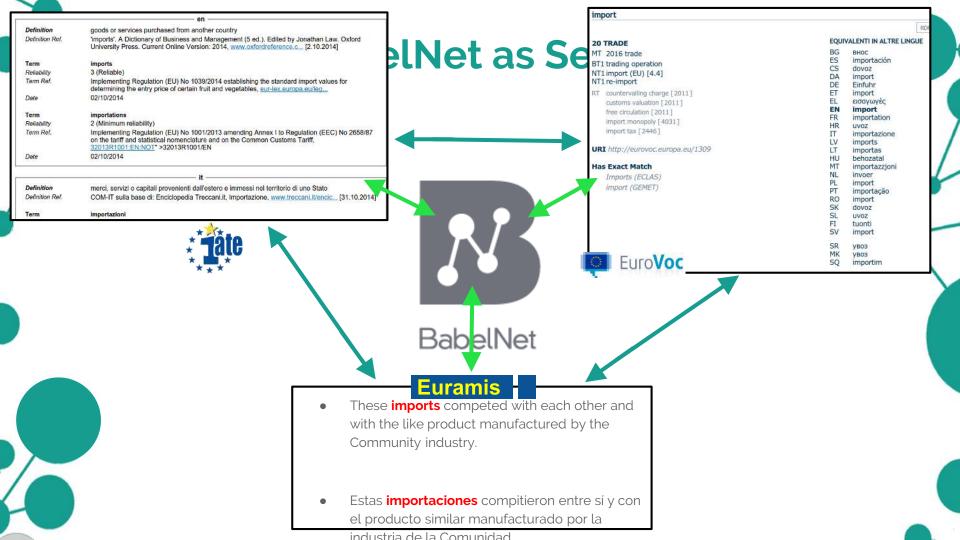
import EQUIVALENTI IN ALTRE LINGUE внос ES importación BT1 trading operation CS dovoz NT1 import (EU) [4,4] DA import NT1 re-import DE Einfuhr ET RT countervailing charge [2011] import customs valuation [2011] εισαγωγές import free circulation [2011] importation import monopoly [4031] HR UVOZ import tax [2446] IT importazione imports URI http://eurovoc.europa.eu/1309 LT importas behozatal Has Exact Match importazzioni NL invoer Imports (ECLAS) PL import Import (GEMET) PT importação RO import SK dovoz SL UVOZ FI tuonti import SR **УВОЗ** MK **VB03** EuroVoc SO importim

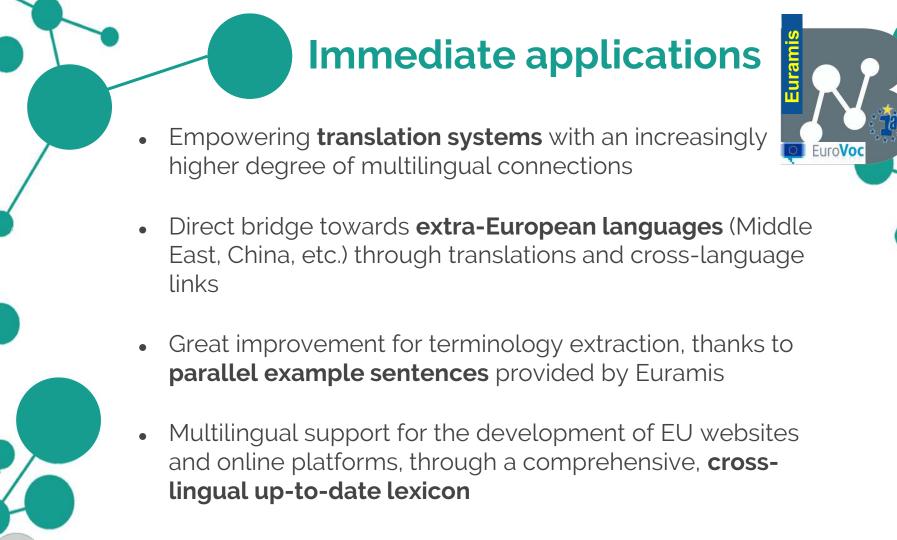
BabelNet

Euramis -

- These imports competed with each other and with the like product manufactured by the Community industry.
- Estas importaciones compitieron entre sí y con el producto similar manufacturado por la

industria de la Comunidad

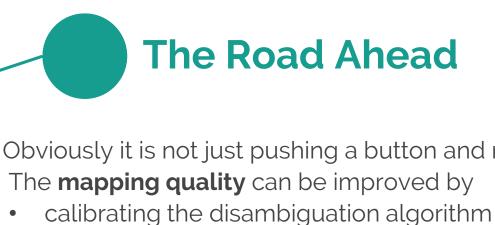








- **Automatic updates** from all the main lexical resources across the web, hence reduced management costs;
- Powerful tool in the hands of EU mediators;
 - Citizens have the right to write to any Community institution or body in one of the languages of the Member States and to receive a response in the same language (Article 24 TFEU, paragraph 4)
- The overall perspective is to integrate through BabelNet many lexical resources, both European and extra-European, using advanced and reliable automatic mapping systems



- Obviously it is not just pushing a button and mapping
- - integrating semantic vector disambiguation techniques
 - and much more
- 2. **Updating** and **syncing** is an issue to be taken into account
 - We are doing it with **BabelNet live!**



REFINE YOUR SEARCH

Only items with links to media

Only With Permission (29)

Image (44)

Text (1)

CAN I USE IT?

MEDIA

Results

à Munich et en Italie | Sandro Botticelli

View at KIK-IRPA, Brussels (Belgium)

Blog ▼

Planche 12 : Esquisse d'après 'La Madone du Magnificat' de Sandro Botticelli Album de 59 planches. Etudes d'après des maîtres anciens exécutées

who:botticelli ×

Home Collections Browse Exhibitions

Add a search term

45 results

™ Image

painting (image-making)



Planche 12 : Esquisse d'après 'La Madonne du Magnificat' de Sandro Botticelli Album de 59 planches. Etudes d'après des maîtres anciens exécutées à Munich et en Italie | Sandro Botticelli

Enable partial matches:

FRENCH

BABELFY!

expanded view | compact view

Planche

12:

Esquisse

d'après 'La

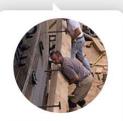
Madonne

du

Magnificat

de

Sandro Botticelli



Planche

A stout length of sawn timber; made in a wide variety of sizes



Esquisse

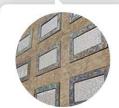
Preliminary drawing for later elaboration

le experience. By using our services, you agree to our use



Madonne

La Vierge à l'Enfant, ou Madone, est un thème récurrent en



Magnificat

(Luke) the canticle of the Virgin Mary (from Luke 1:46 beginning



Sandro Botticelli

Italian painter of mythological and religious paintings





Angyal a töviskoronával | Giovanni Lorenzo Bernini, Giovanni Lorenzo Bernini, Gian Lorenzo Bernini, Gian Lorenzo Bernini, Giovanni Lorenzo Bernini, Lorenzo Bernini, Gian Lorenzo Berni Lorenzo Bernini, Gian Lorenzo Bernini, Gian Lorenzo Bernini, Giovanni Lorenzo Bernini, Le Bernini, Gian Lorenzo Bernini, Gian Lorenz Gian Lorenzo Bernini, and Gian Lorenzo Bernini

Enable partial matches:

AGNOSTIC

BABELFY!

expanded view | compact view

Angyal

a töviskoronával | Giovanni Lorenzo Bernini, Giovanni Lorenzo Bernini ,

Gian Lorenzo Bernini,



Angyal Spiritual being attendant upon God



Lorenzo

he experience. By using our services, you agree to our use of co Italian sculptor and



Gian Lorenzo Bernini

Italian sculptor and